

High-Fidelity Specular SVBRDF Acquisition from Flash Photographs

Michael Tetzlaff

Abstract—Obtaining accurate SVBRDFs from 2D photographs of shiny, heterogeneous 3D objects is a highly sought-after goal for domains like cultural heritage archiving, where it is critical to document color appearance in high fidelity. In prior work such as the promising framework by Nam et al. [1], the problem is simplified by assuming that specular highlights exhibit symmetry and isotropy about an estimated surface normal. The present work builds on this foundation with several significant modifications. Recognizing the importance of the surface normal as an axis of symmetry, we compare nonlinear optimization for normals with a linear approximation proposed by Nam et al. and find that nonlinear optimization is superior to the linear approximation, while noting that the surface normal estimates generally have a very significant impact on the reconstructed color appearance of the object. We also examine the use of a monotonicity constraint for reflectance and develop a generalization that also enforces continuity and smoothness when optimizing continuous monotonic functions like a microfacet distribution. Finally, we explore the impact of simplifying from an arbitrary 1D basis function to a traditional parametric microfacet distribution (GGX), and we find this to be a reasonable approximation that trades some fidelity for practicality in certain applications. Both representations can be used in existing rendering architectures like game engines or online 3D viewers, while retaining accurate color appearance for fidelity-critical applications like cultural heritage or online sales.

Index Terms—SVBRDF acquisition, normal map refinement, non-linear optimization, computational photography, flash photography, photogrammetry, image-based relighting, real-time rendering.

1 INTRODUCTION

SHINY three-dimensional objects, with complex surface appearance characteristics that are sensitive to illumination and viewpoint, can be challenging subjects for digitization and archiving. A single photograph is often not enough to capture the entirety of an object’s color appearance. This has motivated a movement towards 3D digitization within the cultural heritage sector, with many institutions trying to archive fully 3D digital representations of the objects in their collections. Unfortunately, this same class of objects for which 3D digitization would be most valuable exhibit material properties – bumpy surfaces and heterogeneous reflectance – that have been difficult to capture and portray realistically within the domain of computer graphics using established modeling and rendering techniques.

Image-based rendering and relighting refers to the process of synthesizing images based on photographs of a real-world 3D object (often combined with a geometric proxy). Much of the existing work in this area falls into one of two categories. On the one hand, some methods have required photographic hardware that has been prohibitive in cost for all except the most advanced research laboratories, the largest film and television studios, or the most well-funded cultural heritage institutions. On the other hand, other approaches have relied on machine learning to synthesize images using just a single source photograph, relying on a trained model to fill in the missing information. However, a few recent studies have used an approach somewhere in the middle: a sequence of photographs taken using an

ordinary camera with a mounted flash for backscattering illumination [1], [2]. This “middle-ground” approach allows for greater assurances of appearance reproduction fidelity while remaining practical for smaller institutions.

One of these “middle-ground” methods, by Nam et al. [1], uses a *basis function* approach, where spatially-varying reflectance is approximated as a linear combination of a few global isotropic BRDFs with spatially varying weights. In this work, we explore several modifications to the foundational algorithm used by Nam et al. While they addressed both geometry and reflectance recovery, we simplify our implementation by redefining the problem to assume sufficiently accurate macroscale geometry acquired from off-the-shelf photogrammetry software, and focus only on reflectance (microscale) and normal (mesoscale) estimation. Although Nam et al. was novel for its simultaneous acquisition of reflectance and geometry, we believe their approach also has merit for the case considered here where geometry can be obtained through other means with sufficient accuracy.

With this simplification in mind, we consider three modifications to Nam et al.’s implementation. First, Nam et al. make an assumption of isotropy about the surface normal, a symmetry constraint that implicitly assumes accurate normal vectors in order for the axis of symmetry to be correct. In general, the appearance model being optimized is non-linear with respect to surface normals; however, Nam et al. approximate reflectance as constant with respect to normal direction in order to simplify normal estimation to a linear optimization. We study the impact of this simplification by applying non-linear optimization to the full model and contrast it with the results of the linear method.

Second, we observe that the monotonicity constraint used by Nam et al. is equivalent to fitting the basis functions to a linear combination of step functions. To improve the

• M. Tetzlaff is with the Department of Mathematics, Statistics, and Computer Science, University of Wisconsin – Stout, Menomonie, WI, 54751.
E-mail: tetzlaffm@uwstout.edu

Manuscript received TBD; revised TBD.

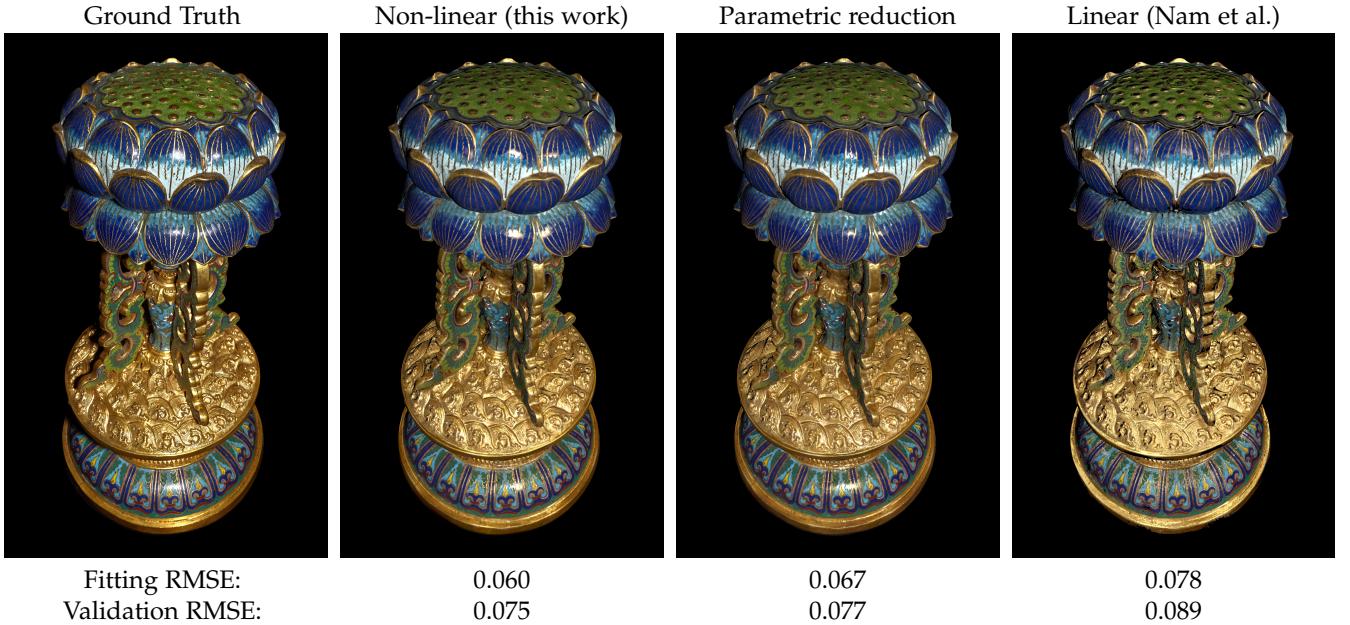


Fig. 1. SVBRDF acquisition for a *cloisonné* altarpiece from 18th century China (94 photographs), in the collection of Mia. From left to right: a ground truth photograph not part of the input dataset, the same viewpoint reconstructed with a technique leveraging nonlinear optimization of mesoscale normals, the same viewpoint after reduction to a parametric model (GGX), the same viewpoint reconstructed with a simplified technique applying linear optimization of mesoscale surface normals. Two error metrics are reported for each case: the fitting RMSE averaged over all the input images, and the validation RMSE averaged over 94 additional images held out for validation. The reproduction quality and RMSE are superior when using nonlinear optimization, and the visual appearance is acceptable even with a reduction to the GGX model, despite a slightly higher RMSE.

smoothness of the BRDF and reduce overfitting of normals, we generalize this using smoothstep functions.

Finally, to take advantage of contexts which do not support custom BRDFs, we show how Nam et al.'s non-parametric, basis-function representation can be converted to a parametric GGX model. An example of a situation where this is useful is real-time global illumination as implemented in modern game engines, which typically does not support custom BRDFs due to the need to precompute indirect lighting, and at the same time exhibits less sensitivity to the BRDF than highlights from discrete light sources.

We evaluate each of these variations by using flash photographs to produce digital representations of several shiny and heterogeneous 3D artifacts from the Minneapolis Institute of Art (Mia) with complex mesoscale geometry and rich specular color appearance. We find that each of the first two extensions improves the results, while the third is a simplification with a reasonably small impact on visual fidelity. One key finding is that surface normals have a very significant impact on the visual quality of the BRDFs generated under constraints of symmetry and isotropy, particularly for objects with complex mesoscale geometry.

The color appearance accuracy achieved by this work is particularly relevant for cultural heritage, where it is expected that any digital depiction must be faithful to the original artifact. It is also highly applicable in online sales and marketing, where the appearance of a digital representation of a product needs to match the physical item being sold. Furthermore, we apply this technique using the Unity game engine* to demonstrate its immediate applicability in an existing commercial rendering architecture.

2 RELATED WORK

2.1 Image-based rendering

Historically, there have been two broad approaches for acquiring reflectance from photographs. The first is image-based rendering, which directly utilizes photographs in the rendering pipeline. The challenge with this paradigm is transferring the color appearance data in the photographs to novel viewpoint and illumination conditions, particularly specular highlights and reflections. Early approaches used highlight warping to improve image quality with a sparse set of images [3]. Other techniques constrain the problem by using radial basis functions [4], [5], compressed sensing [6], or the Kernel Nyström method [7]. A common assumption made by many image-based rendering and relighting methods is that the specular reflectance lobe is both symmetric and isotropic with respect to a “specular normal” direction, which allows the 4D BRDF to be reduced to a 1D function. However, for this assumption to be valid, the specular normal directions must be chosen carefully, a principle that will be explored further in the present work.

Recently, there has been interest in employing neural networks to improve reconstruction quality from a sparse collection of images [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]. Many techniques require only a single photograph, relying on a learned model to fill in the rest. While such approaches have produced impressive results, they are known to be dependent on the quality of the training data, and it is unclear whether state-of-the-art trained models are sufficient for fidelity-sensitive application domains like cultural heritage preservation and archiving. Additionally, while real-time neural network evaluation is viable [24], [25], alternatives that are more lightweight do subjectively have a practical appeal.

*<https://unity.com/>

2.2 Reflectance parameter estimation from photos

The alternative to image-based rendering is to explicitly estimate reflectance parameters from photographs for use in a traditional graphics pipeline. Often this problem is simplified by assuming that the subject is 2D [4], [26], [27], [28], [29], [30], [31], homogeneous [32], [33], or has geometry that is known in advance with high accuracy [34]. Several works have estimated not only the reflectance but also the mesoscale surface geometry, often using iterative least squares optimization for the surface normal estimate [35], [36], [37]. Even when an object is captured under passive illumination, sharp features in the passive illumination conditions can be exploited to extract the width and shape of the reflectance lobe [38], [39], [40], [41], or neural networks may be used to infer these parameters [42], [43]. Polarization has also been used to estimate both the specular reflectance parameters and the mesoscale surface orientation [44], [45].

Specular reflectance is generally a non-linear function, so it is therefore non-trivial to optimize specular reflectance parameters. Techniques that have been used range from Levenberg–Marquardt optimization [35] to brute force linear search in the parameter space [34], [37]. Differentiable rendering, like the pipeline introduced by Laine et al. [46], presents another interesting avenue that could be explored to solve this problem. However, it is also common for a heuristic to be employed that does not guarantee optimality for all possible reflectance lobes, but does still converge to the correct answer when there is a perfect fit for the true reflectance distribution under the assumed parametric model [26], [33], [39]. This latter approach is often good enough even when the true reflectance lobe does not match the expected parametric model.

2.3 Use of photographs with backscattering lighting

Another way to classify methods for reflectance recovery is by the photographic technique used; in particular, the scheme for sampling the viewpoint and lighting. The present work makes use of *backscattering* photography, where the camera and light source are nearly co-located and move together. This configuration has been widely used for acquisition of either flat samples or 3D scenes with limited viewpoint [9], [10], [14], [27], [29], [30], [31], [47], and recently by Tetzlaff and Meyer [2] and Nam et al. [1] to relight arbitrary viewpoints of 3D objects with non-parametric spatially-varying specular reflectance. The recent neural network technique by Gao et al. [20] also used backscattering flash illumination for its input photographs.

The method by Nam et al. [1] is notable as the first technique that is capable of acquiring an SVBRDF from backscattering flash photographs without any loss of generality in the shape of the 3D geometry, the 1D profile of the specular highlights, or the spatially-varying texture. The only significant limiting assumption they make is that of isotropy about the estimated surface normal. Their work was also novel in its ability to acquire the geometry itself as part of the process. The recent work by Bi et al. [48], [49] builds on Nam et al. by improving the method in which the geometry is reconstructed, using discrete volumetric or implicit continuous representations of the scene geometry without an explicit triangulated mesh. In contrast, the focus

of the present work is to derive a technique from Nam et al. that is optimized for use cases where highly accurate 3D models can be reconstructed using commercial photogrammetry software (a typical workflow for cultural heritage institutions). In such cases, only the reflectance and mesoscale geometry (a normal map) need to be recovered, so we only use Nam et al.'s method for recovering reflectance and normals while assuming accurate macroscale geometry as an input to the technique.

3 METHOD

The focus of this work is the development and analysis of several modifications to the innovative SVBRDF acquisition framework proposed by Nam et al. [1]. We begin by defining a baseline implementation and then proceed to describe these variations that will be analyzed to evaluate their impact on image reproduction fidelity.

3.1 Photography

The photographic configuration used for data acquisition is as follows. We use multiple (50 to 500) photographs of an object taken under a backscattering punctual light source, mounted on the camera and nearly co-located with the film. (This is a trivial setup that is even available with the flash on a typical smartphone.) The light source should emit “white” light with a spectral composition that will be acceptable when rendering synthetic images. The images are saved in the sRGB color space and may have non-linear tonemapping applied; thus, prior to other calculations, we will convert colors to a linear space using a calibration chart [50] which provides reference values that can be piecewise interpolated using a standard gamma curve in between these values.

We require camera registration (the position and orientation of the camera when each photo was taken) and a 3D model of the object as input to the rendering pipeline. In this work, a commercial software package called Agisoft Metashape (formerly PhotoScan)[†] was used to calculate this information from the photographs themselves; however, we are not coupled to a particular implementation for this pre-processing step. Using these camera registrations, projective texture mapping [51], [52] is used to map the photographs onto the 3D model, projected and rasterized in the model's 2D texture coordinate space. To account for this projection, in all optimization calculations, we will weight each pixel by the cosine of the angle between the viewing direction ω_r and the normal vector of the polygonal geometry. This compensates for the fact that in the original images, the surface was viewed at a lower effective resolution as the viewing direction becomes closer to parallel with the surface.

Although the light source is very close to the camera, it is still important to account for the displacement of the light from the camera. A straightforward way to accomplish this is by completing a one-time manual calibration that exploits shadows present in the input imagery. In this step, the user views the object from the position of the light source and is instructed to move the viewpoint until all of the shadows are occluded by the object's geometry. At this point, the viewpoint will represent the true position of the light source.

[†]<http://www.agisoft.com/>

3.2 Reflectance Model

The specular reflectance model used throughout this work is a microfacet model after Cook and Torrance [53]. Using the notation that ω_i is the illumination direction with incident radiance L_i , and ω_r is the reflection direction with exitant radiance L_r , the Cook-Torrance BRDF is defined as follows:

$$f(\omega_i, \omega_r) = \frac{L_r}{L_i \cos \theta_i} = \frac{D(\omega_h) G(\omega_i, \omega_r) F(\omega_i, \omega_r)}{4 \cos \theta_i \cos \theta_r} \quad (1)$$

Throughout this work, we use the standard notation that for a direction ω , $\cos \theta$ is the inner product of ω and the normalized normal vector \mathbf{n} . We use the typical notation that $F(\omega_i, \omega_r)$ is the Fresnel reflectance, $G(\omega_i, \omega_r)$ is the microfacet masking and shadowing function, and $D(\omega_h)$ is the distribution of microfacets, defined as a function of the *halfway direction* ω_h that bisects the illumination and reflection directions. Within the context of backscattering photographs, it can be assumed that $F(\omega_i, \omega_r)$ is a constant denoted as F_0 [54].

After Nam et al. [1] in addition to a wide range of prior computer graphics literature, we make the simplification of representing reflectance in an RGB color space rather than treating it as a spectral quantity. For a virtual light source that is also white with the same spectral distribution as the physical light source used to capture the photographs, this approximation will produce accurate results in the same RGB color space. Colored light can be emulated using the simplification that is ubiquitous in computer graphics of multiplying the RGB-valued BRDF by an RGB light color; in most cases this should produce satisfactory results, although physical accuracy can no longer be guaranteed under such conditions. Multispectral BRDF acquisition is not addressed in this work, but generalization to multispectral is possible and could be a direction of future work.

3.3 Baseline Method for Reflectance Estimation

Our objective is to estimate the distribution of microfacets $D(\omega_h)$, assumed to be isotropic around the surface normal. While parameterized formally on the halfway angle θ_h , this distribution is also spatially varying, defined at each sampled point on the object's surface. To solve this problem, we use the framework conceived by Nam et al. [1].

A brief summary of this architecture now follows. The distribution of microfacets is an arbitrary 1D function that is constrained to be monotonic and non-negative, and in practice is estimated and evaluated using a table of discrete values that will be interpolated to form a continuous function. The Fresnel reflectance F_0 is also taken to be absorbed into the microfacet distribution as represented by this lookup table; in other words, the values in this table actually represent $D(\omega_h)F_0$. At each surface position, the local reflectance is represented by spatially-varying weights modulating a linear combination of a small number of global basis functions. An iterative optimization framework is employed that alternately optimizes basis functions, spatially varying weights, and surface normals; K -mean clustering is used to initialize the spatially varying weights. Although the optimization of normals might be underdetermined if no constraints were imposed on the distribution of

microfacets at each sampled surface position, the use of a small number of global basis functions in combination with the monotonicity constraint on each distribution ensures that normal optimization over a large collection of images is generally a sufficiently constrained problem.

Our baseline method differs in a few ways from the implementation described by Nam et al. First, our problem statement differs from theirs in that we assume that the input geometry is sufficiently detailed. Whereas Nam et al. included an additional step in each iteration to refine the geometry itself, our objective is instead to estimate a normal map to increase the appearance of surface detail without modifying the underlying mesh. There is nothing inherent to the rest of our work that would prevent the mesh from being updated as Nam et al. did; however, we did not find this necessary due to the high geometric quality achieved with commercial photogrammetry software. Since Nam et al. used the refined geometry from the previous iteration as the input to the next iteration, we likewise use the refined normals as the initial estimate for the next iteration.

Second, whereas Nam et al. minimized error between the original and reconstructed *reflectance*, our objective is to minimize error in the reconstructed *radiance*, normalized to account for variation in incident radiance. Put another way, this amounts to an additional factor of $\cos \theta_i$ in our calculations, ensuring that pixels where the surface was illuminated at a strong angle (in other words, where $\cos \theta_i$ is small) do not get unduly amplified in importance. This change is justified by how we have defined our objective, which is to build a model that primarily fulfills two purposes: compression of a large number of photographs to a more compact format with minimal loss in image fidelity, and transformation to a format that can predict lighting conditions not captured in the photographs. Since we are only concerned with estimating physical parameters insofar as they facilitate these goals, our primary error metric is the difference in per-pixel image radiance, rather than differences in reflectance or reflectance parameters.

In addition, we applied a couple of minor changes that were trivial to implement and theoretically improve the accuracy of the results. The first of these is that we produce a final estimate of the diffuse albedo map after specular estimation has been completed by subtracting the reconstructed specular reflectance from each source image and performing a per-texel linear regression over the residual. This ensures that the range of the diffuse color space is not limited by the basis functions selected for specular estimation.

We also use the height-correlated Smith function for microfacet masking and shadowing, which has been shown to be more accurate than the V-cavity model employed by Nam et al. [55]. Since the Smith function depends on the microfacet distribution, we approximate it by substituting the masking / shadowing function for the GGX microfacet distribution model (with roughness estimated using the method described in section 3.6), rather than calculating the masking / shadowing function for the non-parametric microfacet distribution. One ramification of the switch from the V-cavity model to the Smith model is that the masking and shadowing function changes after every iteration of the algorithm as the microfacet distribution is improved. While this might raise concerns about convergence, in practice

the effect of the masking and shadowing function is minor (particularly for backscattering images) and does not seem to affect the algorithm's ability to converge.

3.4 Linear vs. Nonlinear Estimation of Normals

Having established a baseline method for simultaneous estimation of spatially varying reflectance and normals, based on the framework laid out by Nam et al., we now proceed to consider three modifications to this baseline. The first and perhaps most significant of these is the use of non-linear optimization to estimate the surface normal at each texel in the normal map.

As described by Nam et al., the cost function being minimized, over K views and P samples per view with sample visibility $v_{p,k}$ (treating the BRDF as spatially varying over \mathbf{x}) is:

$$\sum_{k=1}^K \sum_{p=1}^P v_{p,k} \left([L_r]_{p,k} - f([\omega_i]_{p,k}, [\omega_r]_{p,k}, \mathbf{x}_p, \mathbf{n}_p) [L_i]_{p,k} (\mathbf{n}_p \cdot [\omega_i]_{p,k}) \right)^2 \quad (2)$$

Since f is generally a nonlinear function, this requires nonlinear optimization to estimate \mathbf{n}_p at each surface position p . To simplify the problem, Nam et al. assumed f to be locally constant with respect to \mathbf{n}_p (using the value of \mathbf{n}_p from the previous iteration to evaluate f), reducing the problem to a linear optimization:

$$\sum_{k=1}^K \sum_{p=1}^P v_{p,k} \left([L_r]_{p,k} - f_{p,k} [L_i]_{p,k} (\mathbf{n}_p \cdot [\omega_i]_{p,k}) \right)^2 \quad (3)$$

In contrast, we explore the optimization of the full nonlinear model represented by the cost function in Equation 2, using the Levenberg-Marquardt algorithm [56], [57]. Recognizing this as a nonlinear least squares problem (with L_r as the observations and $L_i f \cos \theta_i$ as the model being fit), the gradient of the cost function is determined by the gradient of the model with respect to the normal vector (the Jacobian). To simplify calculation, we use a coordinate space where only two components of the normal vector $\mathbf{n} = (n_u, n_v, n_w)$ have a nonzero gradient. The details of this change of coordinates are described in an appendix. Applying Equation 1, the gradient of the model comes out to (omitting the (p, k) subscripts for clarity):

$$\begin{aligned} \frac{\partial(f \cos \theta_i)}{\partial n_u} &= f_D(\omega_i \cdot \mathbf{u}) + \frac{G}{4 \cos \theta_r} h_u \frac{\Delta(DF_0)}{\Delta \cos \theta_h} \\ &\quad + \frac{DF_0}{4 \cos^2 \theta_r} \left(\cos \theta_r \frac{\partial G}{\partial n_u} - G(\omega_r \cdot \mathbf{u}) \right) \end{aligned} \quad (4)$$

The gradient with respect to n_v is symmetric, while the gradient with respect to n_w is zero, as noted above. The gradient of G can be found analytically. The full derivation of this gradient (including the gradient of G) is also described in the appendix. In principle, a differentiable rendering framework like Laine et al.'s [46] could be applied instead, but we found the analytic formula straightforward enough to calculate, which has the practical advantage of avoiding additional software dependencies.

The first time we attempt to estimate surface normals, we use the 3D model's vertex normal as our initial estimate. Afterwards, we use the previously estimated normal map as the new initial estimate each time we begin the Levenberg-Marquardt algorithm. Every time the normal estimation step begins (after updating the basis functions and spatially varying weights), we reset the damping factor to 1.0 for all texels. The surface normals for all texels are refined in parallel on the GPU, updating the damping factor independently for each texel after each Levenberg-Marquardt iteration. For every texel, we treat each RGB color channel as a different sample. We determine that the algorithm has converged when either the total root-mean-squared error (RMSE) in reconstruction (across all texels) fails to decrease after a particular iteration, or the error decreases by less than some minimum threshold (we used a threshold of 0.00001 RMSE in cosine-weighted reflectance) for 8 consecutive iterations.

3.5 Arbitrary vs. Smooth Microfacet Distributions

The second extension examines the constraints applied when estimating the global basis functions for representing the distribution of microfacets. The original method by Nam et al. applied a monotonicity constraint when estimating these functions to ensure that they always decrease with angle from the mesoscale surface normal. Numerically, this is essentially equivalent to fitting the distribution of microfacets as a linear combination of step functions with non-negative weights: each step function corresponding to an element in the discrete definition of the microfacet distribution. In practice, we observed that this often leads to discontinuities in the estimation of the microfacet distribution as the algorithm converges to a local minimum (due to ambiguity between reflectance and surface normal).

To address this problem, we generalize to the "smooth-step" function which is common in computer graphics [58]:

$$\text{smoothstep}(t) = 3t^2 - 2t^3 \quad (5)$$

t is typically defined in terms of two edges, but in our implementation, we parameterize it using a single edge (effectively using zero as the second edge).

$$t = \frac{\text{edge} - x}{\text{edge}} \quad (6)$$

Practically speaking, this constraint means the "transition range" of the function is smaller when the edge is close to 0 (allowing for very sharp highlights on highly specular surfaces) but gets larger as the edge increases, preventing the discontinuities seen with only a monotonicity constraint. This ensures that the distribution of microfacets is not only monotonic, but also continuous and smooth, which are also properties typically expected of a microfacet distribution function. Figure 2 illustrates the difference between step and smoothstep functions as they are applied here.

Just as a monotonicity constraint amounts to representing each global basis function as a linear combination of step functions, we can instead substitute these smoothstep functions (using a similar range of edges) and optimize each basis function as a linear combination of smoothsteps. For efficiency, we still ultimately represent each global basis

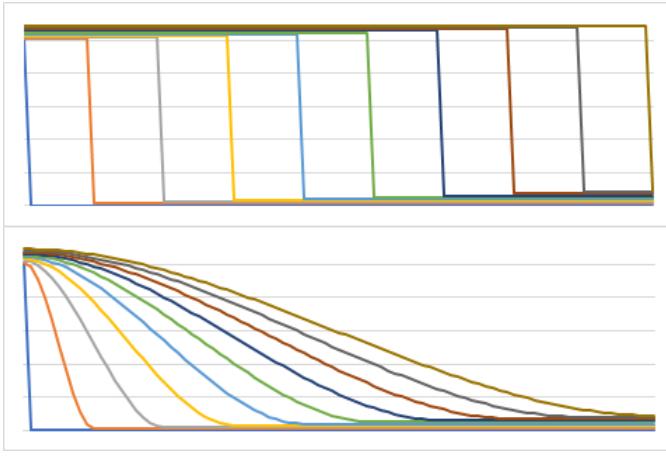


Fig. 2. Top: Representative examples of step functions that can be linearly combined with non-negative weights to form a monotonic function. Bottom: Representative examples of smoothstep functions that can be linearly combined with non-negative weights to form a monotonic function with some smoothness guarantees.

function as a discrete lookup table by precalculating the weighted sum of smoothstep functions for each element of the table. This discrete representation also informs the choices of edges for the range of smoothstep functions; each discrete element becomes an edge for one of the smoothstep functions considered.

Like Nam et al., the global basis functions are optimized simultaneously using linear optimization over all views and surface positions, a single linear system to be solved that provides the optimal solution for each of the global basis functions. (An important clarification must be emphasized here: there are two distinct steps that optimize some linear combination of vectors, which are not to be confused. The first is the optimization of global microfacet distribution functions as a linear combination of smoothsteps, while the second is the optimization of the SVBRDF as a combination of the global functions with spatially-varying weights. For the latter, we use Nam et al.'s method without modification.)

3.6 Estimation of Specular Reflectance Parameters

The final variation considers the use of a parametric analytic model for the distribution of microfacets. The discrete sampling proposed by Nam et al. [1] and employed in the previous sections is generally expected to be superior to a parametric model due to its generality. However, a parametric model is still useful for certain purposes. It can be used to simplify the calculation of the height-correlated Smith masking and shadowing function, which is not strongly observed from backscattering illumination, but can be derived from the distribution of microfacets. It can be used to optimize real-time rendering of environment reflections, which are typically less sensitive to the precise shape of the BRDF than highlights of a punctual light [59]. This calculation will also yield an estimate of the base Fresnel reflectivity F_0 , which can be useful for factoring the Fresnel reflectance from the distribution of microfacets, allowing for simulation of Fresnel reflectance at grazing angles.

For the roughness parameter, the distribution of microfacets is taken to be modeled by the GGX (also called Trowbridge-Reitz) model [60], [61]:

$$D_{\alpha}^{\text{GGX}}(\theta_h) = \frac{\alpha^2}{\pi(1 + (\alpha^2 - 1) \cos^2 \theta_h)^2} \quad (7)$$

We note that while there is a range of excellent prior work on the topic of specular parameter estimation (as discussed in section 2), the approach described here is desirable in its simplicity, particularly given that we already have a refined estimate of the distribution of microfacets and only need to reduce this function to a parametric approximation.

We start by applying the GGX microfacet distribution to the Cook-Torrance model:

$$\frac{L_r}{L_i \cos \theta_i} = \frac{\alpha^2}{\pi(1 + (\alpha^2 - 1) \cos^2 \theta_h)^2} \frac{G(\omega_i, \omega_r) F(\omega_i, \omega_r)}{4 \cos \theta_i \cos \theta_r} \quad (8)$$

For backscattering reflectance, the Fresnel reflectance function reduces to the reflectivity constant F_0 , and to a lesser extent, the microfacet masking and shadowing function is approximately constant for all but the most extreme photographic angles. Applying some algebraic manipulation, we arrive at the following equation:

$$4\pi(1 + (\alpha^2 - 1) \cos^2 \theta_h)^2 \cos \theta_r \frac{L_r}{L_i} \approx \alpha^2 F_0 \quad (9)$$

Nonlinear optimization could be used at this point to find optimal solutions for F_0 and α . However, in practice this has not proven to be necessary; the simpler approach described below has shown itself to be sufficient.

The first step is to evaluate the BRDF where the halfway direction coincides with the normal directions; i.e., due to the backscattering lighting configuration, $\omega_i \approx \omega_r \approx \omega_h = n$. Call the value of the BRDF in this direction f_0 .

We constrain our estimation process by enforcing that f_0 is the magnitude of the parametric BRDF at the specular peak. This establishes a relationship between the surface roughness and the Fresnel reflectivity:

$$4\pi f_0 \alpha^2 = F_0 \quad (10)$$

With this constraint established, each element of the discretely sampled microfacet distribution function that is less than the specular peak can be used to generate an estimate of F_0 :

$$4\pi \cos \theta_r \frac{L_r}{L_i} \left(1 - \cos^2 \theta_h + \frac{F_0}{4\pi f_0} \cos^2 \theta_h \right)^2 = \frac{F_0^2}{4\pi f_0} \quad (11)$$

$$F_0 = 4\pi f_0 \frac{(1 - \cos^2 \theta_h) \sqrt{\cos \theta_r L_r / L_i}}{\sqrt{f_0} - \cos^2 \theta_h \sqrt{\cos \theta_r L_r / L_i}} \quad (12)$$

The estimates derived from all the samples can be averaged to calculate F_0 at each surface position. Since Eq. 12 becomes indeterminate when the denominator is zero, a weighted average is preferable. The weight we use for each sample k is the square of the denominator in Equation 12:

$$w_k = \left(\sqrt{f_0} - \cos^2 \theta_h^k \sqrt{\cos \theta_r^k \frac{L_r^k}{L_i^k}} \right)^2 \quad (13)$$

Finally, α can be estimated by solving Equation 10.

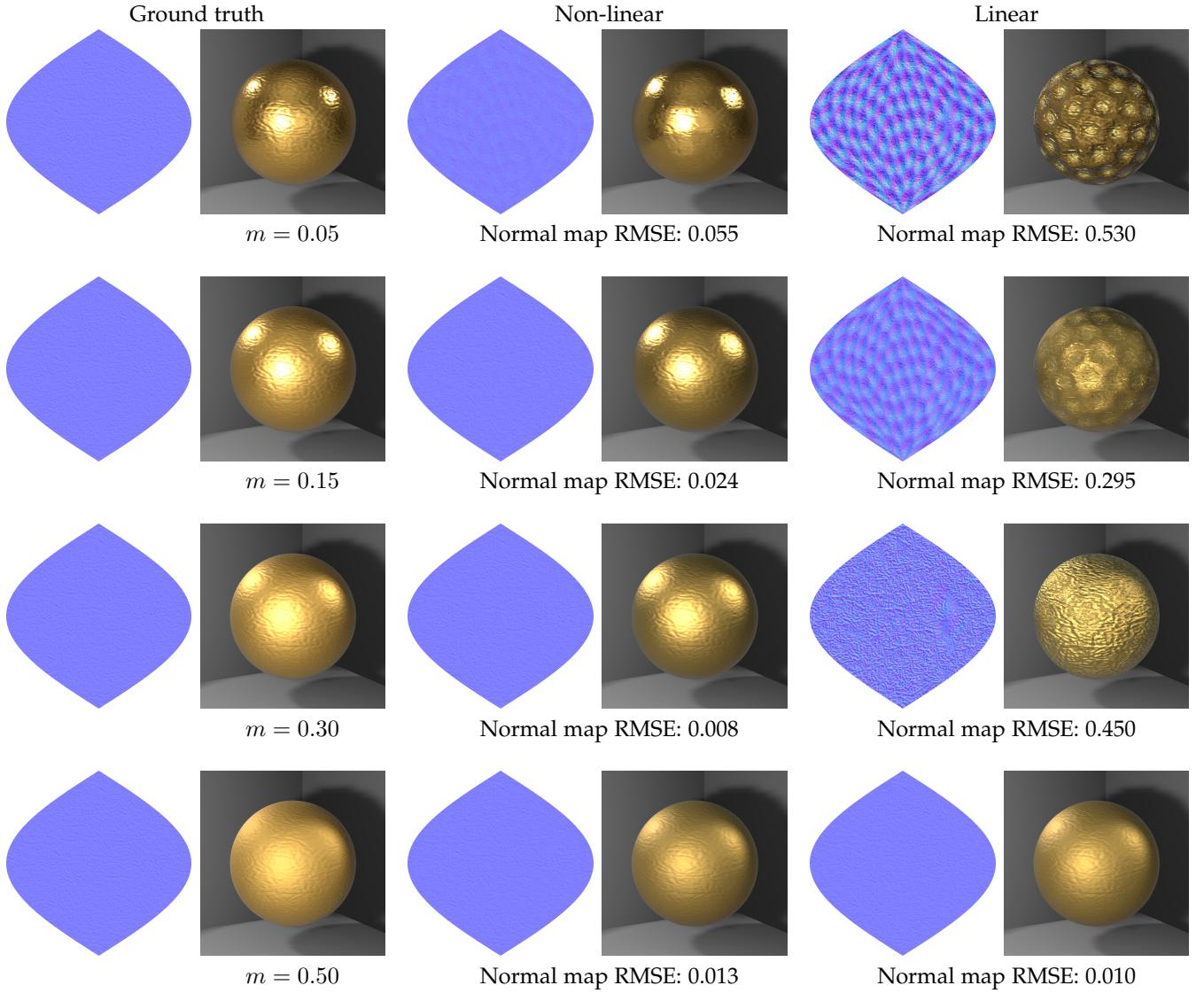


Fig. 3. A ground truth test for normal estimation using synthetic data with varying microscale surface roughness (m). From left to right: The ground truth normal map; the normal map reconstructed from source images using nonlinear optimization (Levenberg-Marquardt); the normal map reconstructed from source images using linear optimization (Nam et al.). (Due to the texture parameterization, the entire normal map was not used.) For each instance (including the ground truth), a rendering of the material in Unity is depicted under novel lighting with respect to the input images. The linear simplification results in severely biased normals, particularly for shinier surfaces. The RMSE between the ground truth normal map and the reconstructed normal map (over the region covered by the texture parameterization) is reported under each reconstruction.

4 RESULTS

4.1 Validation of normal estimation with synthetic data

To evaluate the impact of nonlinear optimization of surface normals as discussed in section 3.4, we begin with an experiment on synthetic images of a simple sphere with a normal map applied, using a metallic Cook-Torrance material with a GGX microfacet distribution. The lighting in this experiment was ideal: an infinitely small point light source perfectly co-located with the camera. Only one basis function was used for this controlled experiment with a homogeneous specular material, and the smoothstep extension of section 3.5 was also applied. To evaluate the effectiveness of each normal optimization method, the estimated normal map can be compared to the one used to generate the source images.

Figure 3 shows the results of this test across four GGX roughness values. The original normal map is shown for

each case alongside the estimated normal maps for both the nonlinear and linear methods. Alongside the normal maps, a rendering of each version in Unity under novel lighting is shown (using the method of section 3.6 to estimate a parametric GGX model for reflections of the environment with Unity's built-in reflection probe system). For each estimate, the RMSE between the ground truth normal map and the reconstruction (across all texels used in the UV mapping) is listed. Whereas the nonlinear algorithm produces reasonable results for most roughness values, just starting to encounter numerical issues at very low roughness, the linear method exhibits severe convergence issues for all but the highest roughness value considered. This phenomenon is analyzed further in section 5.1.

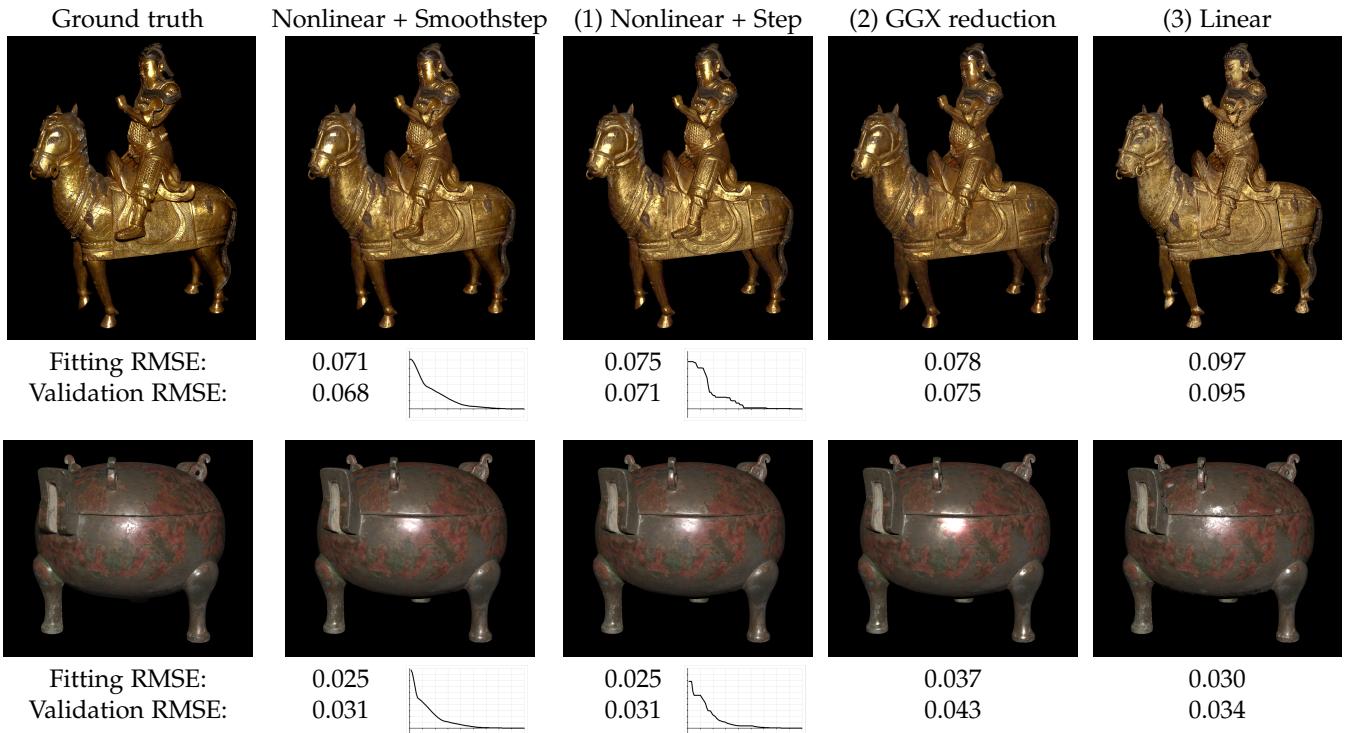


Fig. 4. Top: bronze statue of *Guan Yu* (56 images). Bottom: bronze *ding* food vessel (103 images). From left to right: a ground truth photograph, *not* one of those selected as a source image; the same viewpoint reconstructed using nonlinear optimization of mesoscale surface normals and smoothstep functions for reflectance (one of eight optimized basis functions plotted in the lower right); then, the same viewpoint reconstructed using each of three alternative methods for comparison: (1) using step instead of smoothstep for reflectance, (2) reducing to a parametric model (GGX, using smoothstep for the intermediate representation), and (3) using linear optimization for mesoscale normals. Two error metrics are reported for each case: the fitting RMSE averaged over all the input images, and the validation RMSE averaged over the remaining held-out images.

4.2 Validation of reflectance for real-world objects

Next, we studied three objects from Mia that had been previously photographed from many viewpoints under backscattering flash illumination. The light source used for each object was a conventional spotlight for professional photography, which is reasonably close to the ideal of a point light source. The objects studied include a bronze statue of *Guan Yu* from 16th century China (240 images total, 56 used for reflectance recovery), a *cloisonné* altarpiece from 18th century China (188 images total, 96 used for reflectance recovery), and a bronze *ding* food vessel from 3rd century China (500 images total, 103 used for reflectance recovery). The *Guan Yu* artifact was chosen because of its moderately high degree of metallic specularity and its complex mesoscale geometry. Similarly, the *cloisonné* altarpiece not only has spatially varying texture and geometry, but also exhibits spatially-varying specular characteristics. Both of these objects exemplify use cases for image-based methods that cannot be easily solved with traditional modeling techniques. In addition, the *ding* vessel was chosen to evaluate the technique for an object with high specularity, which would require a very high number of photographs to be sampled sufficiently without the use of basis functions.

Figures 1 and 4 show the results of nonlinear optimization of normals with smoothstep for reflectance (the combination of sections 3.4 and 3.5) against three alternatives: the use of step rather than smoothstep (Figure 4 only), the use of the parametric reduction (section 3.6) and the use of linear optimization for normals. Each example has eight basis functions, and the resolution of all texture maps (diffuse,

normal, specular, and basis weights) is 2048x2048. For each test case, a representative example of a reconstructed validation image is shown along with RMSE metrics, calculated in terms of gamma-corrected image radiance ($\gamma^{-1} = 1/2.2$ for encoding), normalized against incident surface radiance (or equivalently: gamma-corrected, cosine-weighted reflectance). Note that the optimization itself is performed in terms of non-gamma-corrected image radiance, so this is not precisely the same error space used for optimization (which is not gamma-corrected for efficiency and simplicity). For each test case, two aggregate RMSE metrics are listed: one measuring the fitting error across the input images used for reflectance recovery, and the other measuring the reconstruction error across the remaining held-out images reserved for validation.

The figures illustrate that the use of nonlinear optimization substantially improves the quality of the reconstructed image over the linear optimization (alternative 1). It is worth noting that the specular appearance of the object degrades at a rate that is not necessarily captured entirely by the RMSE reported, i.e., for *Guan Yu* the linear results appear to be much worse than the nonlinear version.

The parametric fit (alternative 2) generally has a higher RMSE than the corresponding basis function representation, and is not primarily intended to be used for rendering highlights of punctual lights. However, it is still able to produce a rendering which, upon casual inspection, appears to be a reasonable approximation of the original image. This suggests that it should be appropriate for the less fidelity-critical applications considered (calculation of derived func-

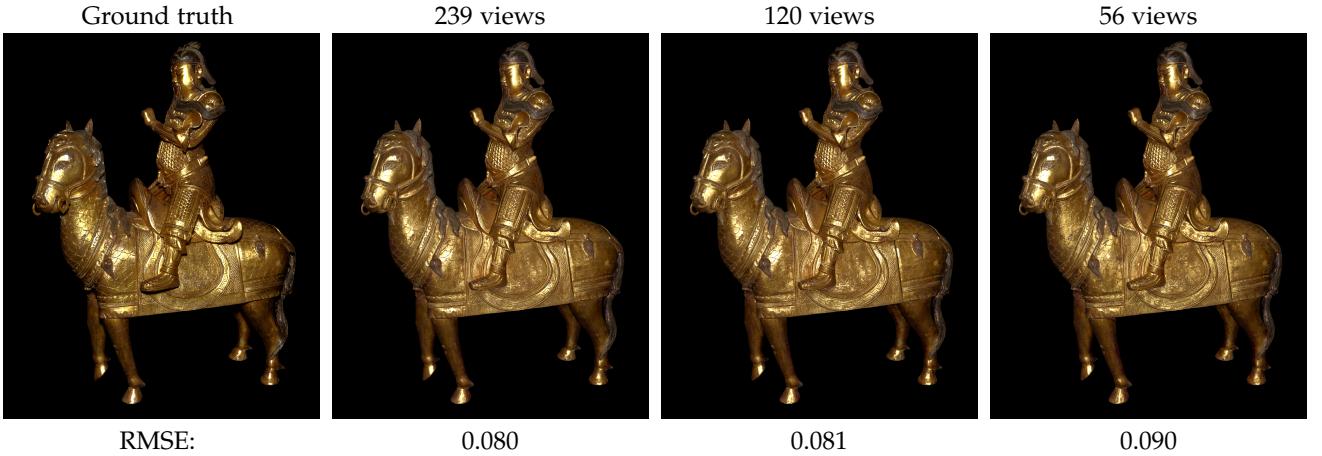


Fig. 5. Results for *Guan Yu* with varying number of input images. From left to right: ground truth, reconstruction from 239 images (all except for the view being reconstructed), reconstruction from 120 images, reconstruction from 56 images. The per-pixel RMSE for the single viewpoint depicted is reported under each reconstructed image.

Artifact	# views	Optimization time
<i>Cloisonné</i> altarpiece	94	44 min.
<i>Cloisonné</i> altarpiece	188	56 min.
<i>ding</i> vessel	103	22 min.
<i>ding</i> vessel	500	3 hr. 36 min.
<i>Guan Yu</i>	56	13 min.
<i>Guan Yu</i>	120	39 min.
<i>Guan Yu</i>	240	1 hr. 18 min.

TABLE 1

Optimization time for several input cases using the nonlinear method for estimation of normals.

tions such as masking / shadowing and Fresnel reflectance, and filtering of environment reflections).

To compare the optimization of basis functions using step (alternative 3) against smoothstep, one of the basis functions is plotted for both the step function results and the original smoothstep results. The discontinuities in the basis functions are very clear when using step functions for both objects. In contrast, the basis functions for smoothstep are more plausible, demonstrating a clear qualitative advantage. How this manifests as a quantitative difference in image quality between step and smoothstep is less pronounced than for the linear / nonlinear comparison, but the smoothstep function does slightly improve reconstruction quality for *Guan Yu*, with no significant difference for the *ding* vessel. While the improvement from smoothstep is probably only observable under punctual light sources like the point light used in Fig. 4, it is worth noting that such illumination is still quite prevalent in real-time applications due to its performance advantage over global illumination. Finally, we note that for the *Guan Yu* example, the use of smoothstep decreased RMSE not only for the validation images, but also for the source images. This suggests that the algorithm may be more likely to converge to a local minimum when using only a monotonicity constraint without smoothness.

For certain objects, like *Guan Yu*, the method improves with the number of input images available. Figure 5 shows the results for *Guan Yu* for three input sets of different sizes, using just a single validation image. While all three produce acceptable results, the image quality and RMSE for

the reconstructed validation image improves as the input size increases.

Table 1 lists the optimization time (using the nonlinear method for estimation of normals) for several input cases. As before, eight basis functions were optimized for each object, using 90 smoothstep functions. The hardware for these benchmarks was a laptop with a six-core Intel i7-8750H CPU, 32 GB of RAM, and an NVIDIA Quadro P3200 GPU with 6 GB of dedicated VRAM. In general, the time complexity for each iteration of the optimization process is linear with respect to the number of input photographs. However, since the optimization process is iterative (and initialized with an element of randomness in the *K*-mean clustering), the number of iterations may also vary based on a range of factors, including not only the number of input images but also the complexity of the surface reflectance and the optimality of the initial *K*-means clustering. For even the slowest case, the optimization time is on the order of hours, using modest hardware, which is on the same order of magnitude as other steps in a photogrammetry pipeline (such as view alignment and point cloud optimization), and is therefore practical for its intended workflow.

Figure 6 shows the objects rendered in the Unity game engine (using the Universal Render Pipeline with a custom shader). For all three objects, nonlinear optimization was utilized for estimating normal maps and smoothstep was used when optimizing the basis functions, with all available photographs taken as source images for optimization (188 photos for the altarpiece, 240 for *Guan Yu*, and 500 for the *ding*). Nam et al.'s basis function representation was used for highlights of punctual light sources, while the parametric GGX model was applied for reflections of the environment (using Unity's built-in reflection probe system). These renderings illustrate the effectiveness and applicability of both representations for the purpose of incorporating museum artifacts into an interactive virtual environment.

5 DISCUSSION

5.1 Analysis of surface normal estimation

The change affecting the optimization results most was the use of nonlinear optimization rather than a linear simplifica-



Fig. 6. Renderings of objects in the Unity game engine. Left: *cloisonné* altarpiece in a virtual museum gallery environment (using imagery of a physical gallery from Mia). Center: *Guan Yu* in a 3D virtual studio environment modeled in Unity. Right: The *ding* vessel using an outdoor environment map and a single directional light. Environment obtained from HDRI Hub. (<https://www.hdri-hub.com/hdrishop/freesamples>)

tion for surface normal estimation. More specifically, strong bias was introduced when assuming the BRDF was constant with respect to a small change in the surface normal (i.e., assuming the reflected radiance to be locally linear with respect to the surface normal). To understand this, it is important to identify the unintended effects that occur with this assumption. First, because normals are optimized every time the weights and basis functions are updated, we must view even the linear simplification as an iterative method. In particular, it is very similar to the Gauss-Newton method for nonlinear optimization, except with an approximate gradient that ignores specularity.

This approximation has two major implications. First, the use of an approximation for the gradient can cause convergence to surface normals that are not optimal. Second, and more significantly, it is known that Gauss-Newton does not guarantee convergence. In practice, our results show that the linear simplification often results in divergence.

The Levenberg-Marquardt algorithm also makes an assumption of local linearity, but is more robust in two respects. First, by calculating the correct gradient, it improves the likelihood of moving the estimate in the right direction. Second, it uses a damping factor that effectively interpolates between Gauss-Newton and the method of gradient descent, which helps to ensure convergence.

One important takeaway from this study is that when evaluating a technique for acquiring an SVBRDF of a surface with unknown mesoscale geometry, it is important to test it on an object with sufficient mesoscale geometric complexity. Prior work (including but not limited to Nam et al.) has often chosen shapes that are mostly smooth for evaluation (even when claiming to support arbitrary 3D geometry), which limits evaluation of the technique's robustness. Objects like *Guan Yu* and the *cloisonné* altarpiece are exemplary in this regard, with complex surface properties that produce a rich surface appearance that is lost if the normals are not reconstructed properly.

It is acknowledged that this work does omit one step from the technique used by Nam et al., that of updating the polygonal geometry. Although we argue that its omission is justified due to the accuracy of photogrammetrically

acquired input geometry, it is possible that this step had a damping effect of its own that prevented their results from diverging to the same extent as what was observed in the present work. Although an effort to reproduce their results with this additional step could reveal whether this was indeed the case, our analysis suggests that it would still be more robust to use nonlinear optimization regardless of whether a geometry update step is included.

5.2 Limitations and Future Work

One concern that is not addressed in this work is the problem of inter-reflections (both diffuse and specular). Gao et al. [20] recently addressed this problem using neural networks. However, the statue of *Guan Yu* used in this work does exhibit minor specular inter-reflections, and our methods still produce results of acceptable quality. Gao et al. showed a problematic case where an object was photographed on top of a specular surface. However, for a controlled studio setting (common in cultural heritage and online sales), both diffuse and specular inter-reflections can be reduced through the practice of photographing the artifact against a black, diffuse background and resting surface. If necessary, an inverse global illumination step might be introduced into the pipeline to factor out diffuse inter-reflection; while this has not yet proved necessary to obtain satisfying results, the implementation of this idea would be an interesting avenue for future study.

Another interesting extension of this work would be to capture more spectral information about the object's reflectance. In principle, the method described in this work could be extended beyond basic red/green/blue color acquisition by putting a spectral filter over the flash light source and taking multiple exposures from the same camera pose with different spectral filters; this would lead to a distinct SVBRDF for each filter used. If the filters were chosen carefully, it might be possible to simulate a spectrally correct rendering of the object under a novel illuminant spectrum by applying a linear combination of these SVBRDFs that best matches the desired illuminant.

6 CONCLUSION

This work refines an architecture introduced by Nam et al. [1] that uses backscattering flash photographs to efficiently recover an SVBRDF for 3D objects that may be shiny and heterogeneous. By considering several modifications to their algorithm, we have made some important observations. First, we have demonstrated the importance of accurate mesoscale surface normals and have shown that nonlinear optimization is the best practice for accurate normal estimation. In contrast, we saw that an algorithm built on assumptions of local linearity without a damping factor can lead to divergence. We have also seen that the use of smoothstep functions may have both qualitative and quantitative advantages over a simple monotonicity constraint for microfacet distribution estimation. Finally, we have suggested that a GGX parametric representation, while slightly inferior to a basis function representation for image reconstruction, may be accurate enough and more efficient for simulating secondary effects like masking / shadowing, Fresnel reflectance, or filtered environment reflections.

Taken together, the findings of this work provide a complete and practical pipeline for accurate acquisition and rendering of color appearance for a wide range of subjects. Unlike light stages and other modes of capturing SVBRDFs, the only photographic hardware that is strictly required is a camera with a mounted or built-in flash, a capability that is even built into most smartphones. At the same time, the method is able to provide more assurance of fidelity than deep learning methods that only use one or two source images and fill in the rest with a pretrained model. Using Nam et al.'s framework with the refinements described in this work, a high degree of color appearance fidelity can be maintained while compressing a model to a representation that can be easily loaded and rendered in a modern game engine or 3D object library, making it immediately applicable for many modern applications.

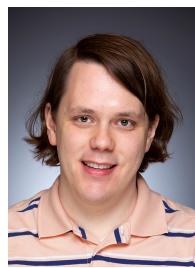
ACKNOWLEDGMENTS

The author would like to thank several individuals for their supporting roles in this work. Gary Meyer, who was very involved in prior research (without which the present work would not exist) provided immensely valuable feedback on an early draft of this work. Dan Dennehy and Charles Walbridge from Mia captured the original flash photographs of the beautiful cultural heritage objects used as examples. Seth Berrier, also from UW-Stout, captured imagery of the gallery environment shown in Fig. 6. Olivia LaValley and Benjamin Odrich both assisted with setting up the test with synthetic images as depicted in Figure 3.

REFERENCES

- [1] G. Nam, J. H. Lee, D. Gutierrez, and M. H. Kim, "Practical svbrdf acquisition of 3d objects with unstructured flash photography," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, pp. 267:1–267:12, 2018.
- [2] M. Tetzlaff and G. Meyer, "Using flash photography and image-based rendering to document cultural heritage artifacts," in *Eurographics Workshop on Graphics and Cultural Heritage*. The Eurographics Association, 2016, pp. 137–146.
- [3] D. N. Wood, D. I. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. H. Salesin, and W. Stuetzle, "Surface light fields for 3d photography," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 287–296.
- [4] J. Filip and R. Vávra, "Fast method of sparse acquisition and reconstruction of view and illumination dependent datasets," *Computers & Graphics*, vol. 37, no. 5, pp. 376–388, 2013.
- [5] T. Zickler, R. Ramamoorthi, S. Enrique, and P. N. Belhumeur, "Reflectance sharing: Predicting appearance from a sparse set of images of a known shape," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1287–1302, 2006.
- [6] P. Sen and S. Darabi, "Compressive dual photography," in *Computer Graphics Forum*, vol. 28, no. 2. Wiley Online Library, 2009, pp. 609–618.
- [7] J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo, "Kernel nyström method for light transport," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 3, pp. 29:1–29:10, 2009.
- [8] P. Ren, Y. Dong, S. Lin, X. Tong, and B. Guo, "Image based relighting using neural networks," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, pp. 111:1–111:12, 2015.
- [9] M. Aittala, T. Aila, and J. Lehtinen, "Reflectance modeling by neural texture synthesis," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 4, pp. 65:1–65:13, 2016.
- [10] V. Deschaintre, M. Aittala, F. Durand, G. Drettakis, and A. Boussau, "Single-image svbrdf capture with a rendering-aware deep network," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 128:1–128:15, 2018.
- [11] K. Kang, Z. Chen, J. Wang, K. Zhou, and H. Wu, "Efficient reflectance capture using an autoencoder," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 127:1–127:10, 2018.
- [12] Z. Li, Z. Xu, R. Ramamoorthi, K. Sunkavalli, and M. Chandraker, "Learning to reconstruct shape and spatially-varying reflectance from a single image," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, pp. 269:1–269:11, 2018.
- [13] Z. Xu, K. Sunkavalli, S. Hadap, and R. Ramamoorthi, "Deep image-based relighting from optimal sparse samples," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 126:1–126:13, 2018.
- [14] D. Gao, X. Li, Y. Dong, P. Peers, K. Xu, and X. Tong, "Deep inverse rendering for high-resolution svbrdf estimation from an arbitrary number of images," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, pp. 134:1–134:15, 2019.
- [15] K. Guo, P. Lincoln, P. Davidson, J. Busch, X. Yu, M. Whalen, G. Harvey, S. Orts-Escalano, R. Pandey, J. Dourgarian et al., "The relightables: volumetric performance capture of humans with realistic relighting," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 6, pp. 217:1–217:19, 2019.
- [16] K. Kang, C. Xie, C. He, M. Yi, M. Gu, Z. Chen, K. Zhou, and H. Wu, "Learning efficient illumination multiplexing for joint capture of reflectance and shape," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 6, pp. 165:1–165:12, 2019.
- [17] A. Meka, C. Haene, R. Pandey, M. Zollhöfer, S. Fanello, G. Fyffe, A. Kowdle, X. Yu, J. Busch, J. Dourgarian et al., "Deep reflectance fields: high-quality facial reflectance field inference from color gradient illumination," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, pp. 77:1–77:12, 2019.
- [18] Z. Xu, S. Bi, K. Sunkavalli, S. Hadap, H. Su, and R. Ramamoorthi, "Deep view synthesis from sparse photometric images," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, pp. 76:1–76:13, 2019.
- [19] M. Boss, V. Jampani, K. Kim, H. P. A. Lensch, and J. Kautz, "Two-shot spatially-varying brdf and shape estimation," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, p. 3982–3991.
- [20] D. Gao, G. Chen, Y. Dong, P. Peers, K. Xu, and X. Tong, "Deferred neural lighting: free-viewpoint relighting from unstructured photographs," *ACM Transactions on Graphics (TOG)*, vol. 39, no. 6, pp. 258:1–258:15, 2020.
- [21] V. Deschaintre, Y. Lin, and A. Ghosh, "Deep polarization imaging for 3d shape and svbrdf acquisition," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, p. 15567–15576.
- [22] J. Guo, S. Lai, C. Tao, Y. Cai, L. Wang, Y. Guo, and L.-Q. Yan, "Highlight-aware two-stream network for single-image svbrdf acquisition," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 123:1–123:14, Jul 2021.
- [23] X. Ma, K. Kang, R. Zhu, H. Wu, and K. Zhou, "Free-form scanning of non-planar appearance with neural trace photography," *ACM*

- Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 124:1–124:13, 2021.
- [24] S. J. Garbin, M. Kowalski, M. Johnson, J. Shotton, and J. Valentin, “Fastnerf: High-fidelity neural rendering at 200fps,” in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, 2021, p. 14346–14355.
- [25] P. Hedman, P. P. Srinivasan, B. Mildenhall, J. T. Barron, and P. Debevec, “Baking neural radiance fields for real-time view synthesis,” in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, 2021, p. 5855–5864.
- [26] A. Gardner, C. Tchou, T. Hawkins, and P. Debevec, “Linear light source reflectometry,” *ACM Transactions on Graphics (TOG)*, vol. 22, no. 3, pp. 749–758, 2003.
- [27] J. A. Paterson, D. Claus, and A. W. Fitzgibbon, “Brdf and geometry capture from extended inhomogeneous samples using flash photography,” *Computer Graphics Forum*, vol. 24, no. 3, pp. 383–391, 2005.
- [28] M. Aittala, T. Weyrich, and J. Lehtinen, “Practical svbrdf capture in the frequency domain,” *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, pp. 110:1–110:12, 2013.
- [29] M. Aittala, T. Weyrich, and J. Lehtinen, “Two-shot svbrdf capture for stationary materials,” *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, pp. 110:1–110:13, 2015.
- [30] J. Riviere, P. Peers, and A. Ghosh, “Mobile surface reflectometry,” *Computer Graphics Forum*, vol. 35, no. 1, pp. 191–202, 2015.
- [31] P. Henzler, V. Deschaintre, N. J. Mitra, and T. Ritschel, “Generative modelling of brdf textures from flash images,” *ACM Transactions on Graphics (TOG)*, vol. 40, no. 6, pp. 284:1–284:13, 2021.
- [32] M. M. Bagher, C. Soler, and N. Holzschuch, “Accurate fitting of measured reflectances using a shifted gamma micro-facet distribution,” in *Computer Graphics Forum*, vol. 31, no. 4. Wiley Online Library, 2012, pp. 1509–1518.
- [33] J. Dupuy, E. Heitz, J.-C. Iehl, P. Poulin, and V. Ostromoukhov, “Extracting microfacet-based brdf parameters from arbitrary materials with power iterations,” *Computer Graphics Forum*, vol. 34, no. 4, pp. 21–30, 2015.
- [34] Z. Zhou, G. Chen, Y. Dong, D. Wipf, Y. Yu, J. Snyder, and X. Tong, “Sparse-as-possible svbrdf acquisition,” *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, pp. 189:1–189:12, 2016.
- [35] H. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel, “Image-based reconstruction of spatial appearance and geometric detail,” *ACM Transactions on Graphics (TOG)*, vol. 22, no. 2, pp. 234–257, 2003.
- [36] B. Tunwattanapong, G. Fyffe, P. Graham, J. Busch, X. Yu, A. Ghosh, and P. Debevec, “Acquiring reflectance and shape from continuous spherical harmonic illumination,” *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, pp. 109:1–109:12, 2013.
- [37] G. Fyffe, P. Graham, B. Tunwattanapong, A. Ghosh, and P. Debevec, “Near-instant capture of high-resolution facial geometry and reflectance,” *Computer Graphics Forum*, vol. 35, no. 2, pp. 353–363, 2016.
- [38] K. Nishino, K. Ikeuchi, and Z. Zhang, “Re-rendering from a sparse set of images,” *Department of Computer Science, Drexel University*, 2005.
- [39] G. Palma, M. Callieri, M. Dellepiane, and R. Scopigno, “A statistical method for svbrdf approximation from video sequences in general lighting conditions,” *Computer Graphics Forum*, vol. 31, no. 4, pp. 1491–1500, 2012.
- [40] Y. Dong, G. Chen, P. Peers, J. Zhang, and X. Tong, “Appearance-from-motion: recovering spatially varying surface reflectance under unknown lighting,” *ACM Transactions on Graphics (TOG)*, vol. 33, no. 6, pp. 187:1–187:12, 2014.
- [41] R. Xia, Y. Dong, P. Peers, and X. Tong, “Recovering shape and spatially-varying surface reflectance under unknown illumination,” *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, pp. 187:1–187:12, 2016.
- [42] P. P. Srinivasan, B. Deng, X. Zhang, M. Tancik, B. Mildenhall, and J. T. Barron, “Nerv: Neural reflectance and visibility fields for relighting and view synthesis,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2021, p. 7495–7504.
- [43] X. Zhang, P. P. Srinivasan, B. Deng, P. Debevec, W. T. Freeman, and J. T. Barron, “Nerfactor: neural factorization of shape and reflectance under an unknown illumination,” *ACM Transactions on Graphics*, vol. 40, no. 6, pp. 237:1–237:18, 2021.
- [44] A. Ghosh, T. Chen, P. Peers, C. A. Wilson, and P. Debevec, “Estimating specular roughness and anisotropy from second order spherical gradient illumination,” in *Computer Graphics Forum*, vol. 28, no. 4. Wiley Online Library, 2009, pp. 1161–1170.
- [45] J. Riviere, I. Reshetouski, L. Filipi, and A. Ghosh, “Polarization imaging reflectometry in the wild,” *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 206:1–206:14, 2017.
- [46] S. Laine, J. Hellsten, T. Karras, Y. Seol, J. Lehtinen, and T. Aila, “Modular primitives for high-performance differentiable rendering,” *ACM Transactions on Graphics (TOG)*, vol. 39, no. 6, p. 194:1–194:14, 2020.
- [47] Z. Hui, K. Sunkavalli, J.-Y. Lee, S. Hadap, J. Wang, and A. C. Sankaranarayanan, “Reflectance capture using univariate sampling of brdfs,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5362–5370.
- [48] S. Bi, Z. Xu, K. Sunkavalli, M. Hašan, Y. Hold-Geoffroy, D. Kriegman, and R. Ramamoorthi, “Deep reflectance volumes: relightable reconstructions from multi-view photometric images,” in *European Conference on Computer Vision (ECCV)*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds. Springer International Publishing, 2020, p. 294–311.
- [49] S. Bi, Z. Xu, P. Srinivasan, B. Mildenhall, K. Sunkavalli, M. Hašan, Y. Hold-Geoffroy, D. Kriegman, and R. Ramamoorthi, “Neural reflectance fields for appearance acquisition,” *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, pp. 39:1–39:11, 2020.
- [50] C. S. McCamy, H. Marcus, J. Davidson *et al.*, “A color-rendition chart,” *Journal of Applied Photographic Engineering*, vol. 2, no. 3, pp. 95–99, 1976.
- [51] P. Debevec, Y. Yu, and G. Borshukov, “Efficient view-dependent image-based rendering with projective texture-mapping,” in *Rendering Techniques '98*. Springer, 1998, pp. 105–116.
- [52] C. Everitt, “Projective texture mapping,” *White paper, NVidia Corporation*, 2001.
- [53] R. L. Cook and K. E. Torrance, “A reflectance model for computer graphics,” *ACM Transactions on Graphics (TOG)*, vol. 1, no. 1, pp. 7–24, 1982.
- [54] G. Nam, J. H. Lee, H. Wu, D. Gutierrez, and M. H. Kim, “Simultaneous acquisition of microscale reflectance and normals.” *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, pp. 185:1–185:11, 2016.
- [55] E. Heitz, “Understanding the masking-shadowing function in microfacet-based brdfs,” *Journal of Computer Graphics Techniques*, vol. 3, no. 2, pp. 32–91, 2014.
- [56] K. Levenberg, “A method for the solution of certain non-linear problems in least squares,” *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164–168, 1944.
- [57] D. W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.
- [58] Khronos Group, “smoothstep,” 2014. [Online]. Available: <https://www.khronos.org/registry/OpenGL-Refpages/gl4/html/smoothstep.xhtml>
- [59] M. Ashikhmin and A. Ghosh, “Simple blurry reflections with environment maps,” *Journal of Graphics Tools*, vol. 7, no. 4, pp. 3–8, 2002.
- [60] T. Trowbridge and K. P. Reitz, “Average irregularity representation of a rough surface for ray reflection,” *JOSA*, vol. 65, no. 5, pp. 531–536, 1975.
- [61] B. Walter, S. R. Marschner, H. Li, and K. E. Torrance, “Microfacet models for refraction through rough surfaces,” in *Proceedings of the 18th Eurographics Conference on Rendering Techniques*. Eurographics Association, 2007, pp. 195–206.



Michael Tetzlaff is an Assistant Professor of Computer Science at the University of Wisconsin - Stout. He received his B.S. from Bethel University in St Paul, and his M.S. and Ph.D. from the University of Minnesota. His research interests include interactive image-based computer graphics techniques and their applications in cultural heritage.